



HAL
open science

Interlots:LexInnova/Spécifications/Structure de la base lexicale

Mathieu Mangeot

► **To cite this version:**

Mathieu Mangeot. Interlots:LexInnova/Spécifications/Structure de la base lexicale. [Rapport Technique] Université Grenoble - Alpes. 2016. hal-02165755

HAL Id: hal-02165755

<https://hal.science/hal-02165755>

Submitted on 26 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

Interlots:LexInnova/Spécifications/Structure de la base lexicale

Auteur : Mathieu Mangeot

Date : 9 juin 2016

Macrostructure

En attendant plus d'explications, une macrostructure en volumes a été proposée pour la base lexicale

Cette macrostructure s'appuie sur les questions posées à d'autres lots (et nous mêmes).

Nomenclature

L'unité lexicale est fondée sur l'association d'un lexème et d'une catégorie grammaticale. Il n'y a pas de distinction de vocables homographes. En français, voler (avion) et voler (dérober) seront deux sens de la même entrée "Voler, verbe". Pour le français, on distingue le masculin du féminin pour les noms communs. "horloger" et "horlogère" seront deux articles séparés.

Les locutions, expressions figées, expressions semi-figées sont des entrées à part entière. En français, voici un exemple : "manger les pissenlits par la racine". Ces entrées peuvent être reliées par des liens internes au volume aux entrées des lexèmes composant l'expression (ici manger, pissenlit et racine). Ce choix est motivé pour trois raisons :

- nous nous inspirons de la lexicographie explicative et combinatoire (issue de la théorie sens-texte) qui distingue les locutions comme des entrées à part entière;
- ces locutions peuvent être traduites dans d'autres langues par un seul mot ou par une expression équivalente mais donc aucun mot de l'expression n'est traduction d'un mot de l'autre expression;
- la possibilité d'ajouter des liens internes permet de retrouver les expressions utilisant un mot à partir de ce mot. Par exemple dans l'entrée "pissenlit" se trouvera un lien vers l'entrée "manger les pissenlits par la racine"

Microstructure

La microstructure est en constante évolution. Elle reprend l'état courant des réflexions et les besoins des différents modules utilisant la base lexicale. Nous utiliserons pour l'instant la même microstructure pour toutes les langues et pour toutes les ressources (dictionnaire de référence, lexique institutionnel, lexique de groupe et lexique personnel). D'une langue à l'autre, certains éléments ne seront pas renseignés (ex : translittération pour le français). Certaines listes fermées de valeurs (ex : les catégories grammaticales) dépendent également des langues.

Pour la définition et les collocations, une idée est de reprendre également les pratiques de la lexicographie explicative et combinatoire. La définition est vue comme une formule sémantique et les collocations sont modélisées avec des fonctions lexico-sémantiques. Pour plus d'informations, voir les dictionnaires Papillon et Réseau Lexical du Français.

Dans le lexique personnel, une partie des articles sera réservée aux applications utilisant les lexiques pour qu'elles puissent stocker des informations.

Microstructure actuelle en XML

Exemple d'article en espagnol. Note : les accents dans les balises XML ont été supprimés ici pour des raisons d'esthétique d'affichage (<définition> => <definition>). Ils sont maintenus dans la version officielle.

```
<article id="esp.allí.1">
  <forme>
    <vedette>allí</vedette>
  </forme>
  <classe-gram>adv.</classe-gram>
  <semantique>
    <sens>
      <definition>Indica lugar alejado del que habla y del que escucha. Con verbos
de estado,
      significa `en aquel lugar o sitio'. Con verbos de movimiento,
significa `a aquel lugar':
      no quiero ir allí. Puede ir precedido de las preps. de, desde,
hacia, hasta y por: ¿vienes
      andando desde allí? Precede a otros adverbios: allí arriba;
allí abajo.
    </definition>
    <traduction lang="fra">là</traduction>
    <exemples>
      <exemple>
        <texte-exemple lang="esp">Descansaremos allí.</texte-exemple>
        <traduction-exemple lang="fra">Nous nous reposerons là.</traduction-exemple>
      </exemple>
      <exemple>
        <texte-exemple lang="esp">No quiero ir allí.</texte-exemple>
        <traduction-exemple lang="fra">Je ne veux pas aller là.</traduction-exemple>
      </exemple>
    </exemples>
  </sens>
  <expressions>
  </expressions>
</semantique>
</article>
```

Microstructure prévue à moyen terme

Exemple en espagnol :

```
<article id="esp.allí.1">
  <forme>
    <vedette>allí</vedette>
    <!-- prononciation en Alphabet Phonétique International de préférence -->
    <prononciation>aj'i</prononciation>
    <!-- découpage syllabique au moins pour check your smile -->
    <syllabes>a/llí</syllabes>
  <niveaux>
```

```

<CECLR>A1</CECRL>
  <methode nom="xxx">Leçon 3</methode>
</niveaux>
</forme>
<classe-gram>adv.</classe-gram>
<semantique>
  <sens id="esp.allí.1.1">
    <definition>Indica lugar alejado del que habla y del que escucha. Con verbos
de estado,
      significa `en aquel lugar o sitio'. Con verbos de movimiento,
significa `a aquel lugar':
      no quiero ir allí. Puede ir precedido de las preps. de, desde,
hacia, hasta y por: ¿vienes
      andando desde allí? Precede a otros adverbios: allí arriba;
allí abajo.
    </definition>
    <lien-traduction volume="innovalangues_axi" lang="axi" type="pivot" article="axi.là.xxx" />
  <exemples>
    <exemple>
      <texte-exemple lang="esp">Descansaremos allí.</texte-exemple>
      <traduction-exemple lang="fra">Nous nous reposerons là.</traduction-exemple>
    </exemple>
    <exemple>
      <texte-exemple lang="esp">No quiero ir allí.</texte-exemple>
      <traduction-exemple lang="fra">Je ne veux pas aller là.</traduction-exemple>
    </exemple>
  </exemples>
</sens>
</semantique>
<stockage-traces>
  <application nom="Magic World">.....</application>
</stockage-traces>
</article>

```

Exemple en français :

```

<article id="fra.abandon.1">
  <forme>
    <vedette>abandon</vedette>
    <!-- prononciation en Alphabet Phonétique International de préférence -->
    <prononciation>\a.bã.dõ</prononciation>
    <!-- découpage syllabique au moins pour check your smile -->
    <syllabes>a/ban/don</syllabes>
    <formes-flechies>
      <!-- est-il nécessaire de répéter le lemme ? -->
      <forme-flechie><forme>abandon</forme><nombre>singulier</nombre></forme-flechie>
      <forme-flechie><forme>abandons</forme><nombre>pluriel</nombre></forme-flechie>
    </formes-flechies>
  </forme>
</article>

```

```

    <niveaux>
      <CECLR>A2</CECRL>
      <methode nom="xxx">Leçon 3</methode>
    </niveaux>
  </forme>
  <classe-gram>nom commun masculin</classe-gram>
  <semantique>
    <sens id="fra.abandon.1.1">
      <definition>État d'une personne, d'une chose abandonnée.</definition>
    </sens>
    <sens id="fra.abandon.1.2">
      <definition>Action d'abandonner, de délaisser.</definition>
      <fonctions-lexicales>
        <fonction><nom>QSyn</nom><valeurs>
          <lien type="interne" cible="fra.arrêt.1" sens="fra.arrêt.1.1">arrêt</lien>
          <lien type="interne" cible="fra.cessation.1" sens="xxx">cessation</lien>
          <lien type="interne" cible="fra.suspension.1" sens="xxx">suspension</lien>
          <lien type="interne" cible="fra.abdication.1" sens="xxx">abdication</lien>
        </fonction>
      </fonctions-lexicales>
    </sens>
    <sens id="fra.abandon.1.x">
      <definition>...</definition>
    </sens>
  </semantique>
  <stockage-traces>
    <application nom="Magic World">.....</application>
  </stockage-traces>
</article>

```

Exemple en japonais :

```

<article id="jpn.時計.1">
  <forme>
    <vedette>時計</vedette>
    <lecture>tokei</lecture>
    <ecriture>とけい</ecriture>
    <!-- prononciation en Alphabet Phonétique International de préférence -->
    <prononciation>est-ce nécessaire ou redondant par rapport au romaji ???</prononciation>
    <!-- découpage syllabique au moins pour check your smile -->
    <syllabes>?? est-ce nécessaire ou redondant par rapport au hiragana ?</syllabes>
  <niveaux>
    <CECLR>A1</CECRL>
    <methode nom="Parlons japonais">Leçon 3</methode>
  </niveaux>
</forme>
<classe-gram>名詞</classe-gram>
<semantique>

```

```
<sens>
  <definition></definition>
  <lien-traduction volume="innovalangues_axi" lang="axi" type="pivot" article="axi.horloge.xxx" />
</sens>
</semantique>
<stockage-traces>
  <application nom="Magic World">.....</application>
</stockage-traces>
</article>
```

Conclusion

Il reste un certain nombre de points à décider avant l'implantation effective des structures. Pour les exemples, notamment, il faudra décider si les traductions sont stockées directement dans les articles ou si une axie d'exemple est créée avec des liens vers les articles de chaque langue.

Sources et contributeurs de l'article

Interlots:LexInnova/Spécifications/Structure de la base lexicale *Source:* <http://wiki.innovalangues.net/index.php?oldid=17874> *Contributeurs:* Loizbek, Mangeot

Licence

Creative Commons Attribution-Non Commercial-Share Alike
<https://creativecommons.org/licenses/by-nc-sa/3.0/>
